Journal of Biomedical Informatics 56 (2015) 65–68

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Commentary

The perils of meta-regression to identify clinical decision support system success factors

Christopher L. Fillmore *, Casey A. Rommel, Brandon M. Welch [1], Mingyuan Zhang, Kensaku Kawamoto

Department of Biomedical Informatics, University of Utah, 421 Wakara Way, Ste 140, Salt Lake City, UT 84108, United States

ARTICLE INFO

Article history:
Received 24 November 2014
Revised 12 February 2015
Accepted 11 May 2015
Available online 18 May 2015

Keywords:
Clinical decision support
Meta-regression

ABSTRACT

Clinical decision support interventions are typically heterogeneous in nature, making it difficult to identify why some interventions succeed while others do not. One approach to identify factors important to the success of health information systems is the use of meta-regression techniques, in which potential explanatory factors are correlated with the outcome of interest. This approach, however, can result in misleading conclusions due to several issues. In this manuscript, we present a cautionary case study in the context of clinical decision support systems to illustrate the limitations of this type of analysis. We then discuss implications and recommendations for future work aimed at identifying success factors of medical informatics interventions. In particular, we identify the need for head-to-head trials in which the importance of system features is directly evaluated in a prospective manner.

1. Introduction

Clinical decision support (CDS) interventions are typically heterogeneous in nature, making it difficult to identify why some interventions succeed while others do not [1]. In recent years, investigators have sought to address this difficulty through the use of meta-regression techniques [1–6]. In this approach, investigators apply regression techniques to identify features of information systems (e.g., workflow integration or patient engagement) that are significantly associated with desired outcomes. However, meta-regression analysis has recognized limitations [7–9]. A particularly important limitation is the observational nature of this type of analysis. As such, causal relationships cannot be illuminated by this approach alone. To illustrate why care should be taken when performing or interpreting this type of analysis, this manuscript provides a cautionary case study in the context of CDS evaluation.

In 2013, the *British Medical Journal* published a meta-regression analysis by Roshanov et al. [6] of randomized controlled trials (RCTs) of CDS systems. The study sought to identify features associated with effective systems and resulted in conclusions that differed significantly from similar studies, including a prior meta-regression analysis in the *British Medical Journal* by Kawamoto et al. [1] Notably, Roshanov et al. [6] found that advice given automatically in workflow was not significantly associated with system success in their initial model. As a result, this feature was removed from their final model.

This particular finding was unexpected, as it differed significantly from findings suggested by previous reviews addressing clinical decision support [1,2,4,10–12]. Moreover, RCTs that directly evaluated the importance of this feature have found automatic provision to be important [13,14]. Specifically, within the clinical context of hyperlipidemia management, van Wyk et al. [14] compared alerts provided automatically to physicians within an EHR versus on-demand CDS which had to be proactively accessed by physicians within the same EHR. In this cluster RCT involving 38 Dutch general practices and 87,886 patients, 65% of the patients requiring screening were screened in the automatic CDS group, as compared to 35% in the on-demand CDS group (adjust relative risk 1.40; 95% confidence interval [CI], 1.08 to 1.81) [14]. In another RCT directly evaluating the importance of providing CDS automatically, Scheepers-Hoeks et al. [13] compared alerts provided automatically to physicians within an EHR versus the same information provided on-demand in the EHR. In this RCT, which was conducted in an intensive care unit (ICU) regarding 13 locally developed clinical rules, compliance with the CDS recommendations was 41% in the automatic alerting group, versus 19% in the on-demand EHR group ($p < 0.0001$) [13]. Such findings from head-to-head RCTs must be considered very seriously, as they directly evaluate causative relationships between a CDS feature and system impact, rather than merely correlation as in the case of a meta-regression analysis.

* Corresponding author. Present address: Homer Warner Center for Informatics Research, Intermountain Healthcare, 5171 S Cottonwood St, Ste 220, Salt Lake City, UT 84107, United States.
E-mail address: chris.fillmore@imail.org (C.L. Fillmore).
[1] Biomedical Informatics Center, Medical University of South Carolina, 55 Bee St, Charleston, SC 29425, United States.

http://dx.doi.org/10.1016/j.jbi.2015.05.007
1532-0464/

Given the stark discrepancy in findings between the meta-regression analysis by Roshanov et al. [6] and these prior studies on the topic, we sought to discover an explanation for these findings. Here, we describe our findings and discuss implications for future work aimed at identifying success factors of medical informatics interventions.

2. Methods

2.1. Initial investigation

Based on results from other studies identifying the importance of automatic CDS provision, as well as our initial review of the source data set provided in Appendix of the Roshanov et al. [6] article, we suspected that the differences between the two meta-regression analyses were more likely due to discrepancies in the source data rather than differences in statistical analysis methods. In particular, in the meta-regression analysis by Kawamoto et al. [1], CDS systems that automatically provided their advice had a success rate of 75%, versus a success rate of 0% for systems that did not (difference = 75%). In contrast, in the systematic review by Roshanov et al. [6], the difference in success rate was only 60% versus 54% (difference = 6%).

In examining potential reasons for the differences in the source data set, we found that the differences appeared to stem from discrepancies in the determination of whether specific explanatory features were present or absent in a given CDS system. In particular, Roshanov et al. [6] appeared to consider many CDS systems whose use was required by a study protocol to not be automatically provided as a part of clinician workflow, whereas the end result would be the same: clinicians would always be exposed to the intervention.

2.2. Hypothesis

In our opinion, mandated, protocol-driven use of a CDS system is functionally equivalent to automatic provision of CDS. Therefore, we hypothesized that the results would be more consistent with prior studies if we considered protocol-driven provision of CDS to constitute automatic provision. To test this hypothesis, we repeated the same statistical analysis conducted by Roshanov et al. using this updated definition of automatic CDS provision.

2.3. Dataset

We used the dataset of 162 randomized controlled trials of CDS systems identified by Roshanov et al. [6]

2.4. Outcome and explanatory variables

As in the original systematic review, "effective" systems were defined as those systems that improved primary (or 50% of secondary) reported outcomes of process of care or patient health. For the outcome measure, we maintained the determinations made by Roshanov et al. [6] As in the original study, we focused on six potential explanatory variables for CDS outcomes: automatic provision of CDS, development by authors, feedback at the time of care, advice presented in electronic charting or order entry, advice for patient, and requires reason for override.

2.5. Feature reanalysis

We re-assessed each source study for the presence of "automatic provision of CDS" using an updated definition that included cases where use of the CDS was mandated by protocol. If both the intervention and control groups involved use of a clinical information management tool (e.g., EHR or computerized provider order entry [CPOE] system), the CDS was considered to be automatically provided if exposure to the CDS did not require any end-user initiative beyond simply using the base information management tool. If it was unclear whether the CDS was provided automatically, we deferred to the determination made by Roshanov et al. [6].

In addition, we sought to improve the quality of the data set for the reanalysis by completing incomplete explanatory variable classifications. Specifically, we reanalyzed the "author developed" and "feedback at time of care" features for trials where Roshanov et al. [6] were unable to determine the presence or absence of those features. We considered a CDS intervention to have been developed by the authors if the study stated or implied that one or more of the study authors were involved in the development of the system. If the authors developed the underlying clinical algorithm or knowledge base used in a CDS system, we also considered the CDS intervention to have been developed by the authors. We considered a CDS intervention to have delivered feedback at the time of care if the CDS was provided to the clinician while the clinician was with the patient in question for the purposes of clinical care. If we were unable to make a determination based on the manuscript, we attempted to make direct contact with the authors to make these determinations. If it was unclear whether the CDS system was author developed or provided feedback at the time of care, and if we were unable to make contact with the authors, we maintained the unknown status.

The presence or absence of features was determined through the consensus of four of the authors (CLF, MZ, BMW, and KK). These reviewers consisted of two physician informaticists (KK and CLF), one nurse informaticist (MZ), and one additional health informaticist (BMW).

2.6. Exclusions

We excluded studies where the intervention arm with the CDS system was disadvantaged in a significant way compared to the control arm. Example disadvantages included care delivery by practitioners with significantly less training (e.g., nurse versus physician), or the removal of a key diagnostic resource.

2.7. Statistical analysis

The original data set and statistical analysis file used in the manuscript by Roshanov et al. [6] was kindly provided by the authors. Using this data set and analysis file, the primary analysis from the original study was replicated. Following this replication of findings, the data set was modified as described above, and the primary analysis was repeated with the updated data. In brief, the primary analysis as developed by Roshanov et al. [6], and replicated here without modification, consisted of the following: (i) the selection of explanatory variables for inclusion in the final primary model, with variables selected if they were associated with the study outcomes with a $p < 0.1$; and (ii) the development of a final model to explain the study outcomes using these selected explanatory variables. For these analyses, Firth's profile penalized likelihood method was used as in the original study. All statistical analyses were performed using Stata 13.1 [15].

3. Results

3.1. Feature reanalysis

In the reanalysis, we revised the "automatic provisioning" feature for 46 (28%) of the original 162 studies in the data set. Of these

Table 1
Meta-regression results. Figures are odds ratios (95% confidence interval), *P* value.

Initial model	Original study (148 trials)	Reanalysis (159 trials)
Developed by authors	3.52 (1.34–9.27), 0.01	3.87 (1.55–9.63), 0.003
Advice automatically in workflow	1.48 (0.62–3.52), 0.38	2.92 (0.82–10.38), 0.088
Advice at time of care	0.61 (0.21–1.77), 0.35	0.88 (0.32–2.41), 0.8
Advice presented in electronic charting or order entry	0.33 (0.14–0.76), 0.01	0.37 (0.18–0.79), 0.009
Provides advice for patients	2.54 (0.98–6.57), 0.05	2.58 (1.00–6.66), 0.042
Requires reason for override	10.69 (1.87–61.02), 0.001	7.06 (1.78–28.03), 0.001
Final model	(150 trials)	(159 trials)
Developed by authors	4.35 (1.66–11.44), 0.002	3.96 (1.59–9.86), 0.002
Advice automatically in workflow	N/A	3.03 (0.86–10.68), 0.078
Advice presented in electronic charting or order entry	0.37 (0.17–0.80), 0.01	0.36 (0.17–0.77), 0.007
Provides advice for patients	2.77 (1.07–7.17), 0.03	2.62 (1.02–6.78), 0.038
Requires reason for override	11.23 (1.98–63.72), <0.001	7.25 (1.82–28.89), <0.001

46 studies, 42 were revised from the absence to the presence of the feature.

In the original paper, 14 studies did not have a determination for either the "developed by authors" or "feedback at time of care" features and thus were not included in the original primary analysis. Of these 14 studies with missing data, we were able to complete the data in 13 of the studies.

In total, these revisions were applied to 50 (31%) of the studies in the data set. The Supplemental online Appendix provides details regarding these revisions.

3.2. Exclusions

We excluded two studies where the CDS intervention arm was significantly disadvantaged compared to the control arm. In one study, the performance of nurses using a CDS system was compared to that of physicians who did not use the CDS system [16]. In the other excluded study, in which the outcome of interest was the diagnosis of acute small bowel obstruction, the CDS intervention arm did not have access to contrast radiography, a key diagnostic resource available in the control arm [17].

3.3. Meta-regression analysis results

Table 1 summarizes the findings from the reanalysis. As shown in the table, automatic provision of CDS was associated with study outcome with a $p < 0.1$ in the initial model. Therefore, this feature was eligible for inclusion in the final model. This finding was in contrast to the results from the original study, where the significance of the association was $p = 0.38$, such that this feature was excluded from the final model. The end results for the other features remained the same; "advice at time of care" did not meet inclusion criteria for the final model, while the remaining features were included.

In the final model, "automatic provision of CDS" was found to be nearly statistically significant ($p = 0.078$). The remaining features that were identified as being important in the original study remained so in the final model.

4. Discussion

4.1. Summary of findings

In their original study, Roshanov et al. [6] concluded that automatic provisioning of CDS was not an important factor in determining the effectiveness of CDS systems. In contrast, automatic provisioning of CDS met criteria for inclusion in our final model and approached statistical significance in the final analysis ($p < 0.1$). As noted in the introduction, our findings align with the conclusions from RCTs that have directly evaluated the importance of this feature, which have found that automatic provision of CDS is a critical factor for determining the effectiveness of CDS systems [13,14].

4.2. Implications

Beyond the obvious implications related to automatic provision of CDS, this reanalysis serves to highlight potential limitations of meta-regression techniques. One such limitation is that meta-regression results can hinge upon how explanatory factors are defined and subsequently abstracted. In this reanalysis, we believe our definition of automatic provisioning more accurately reflects the reality of automatically providing CDS within provider workflow. A clear implication, then, is that meta-regression analyses are highly dependent on the exact definitions used for explanatory factors and the reliability of their abstraction.

Another implication of this study is that a meta-regression analysis can benefit from secondary reanalysis by an independent group. While it is assumed that different groups would reach the same conclusions from the same source literature, the only empirical way to validate this assumption is for the reanalysis to actually be conducted. Of note, a confirmatory reanalysis may be difficult to publish, similar to how a negative study can be difficult to publish. Consequently, an important challenge will be identifying how to motivate and reward the conduct of such reanalyses.

Furthermore, the most critical limitation of meta-regression techniques is the inability to evaluate for causative relationships between variables. As such, we believe another vital implication of our findings is that more RCTs that directly evaluate the importance of CDS features are needed. Additionally, given the multi-faceted nature of CDS interventions, researchers should consider the use of modern trial designs such as adaptive and pragmatic trials that may allow for more appropriate evaluations of interventions within complex sociotechnical settings. Considering the inherent challenges of inferring causality or demonstrating real-world effectiveness from meta-regression analyses, direct trial evidence will be invaluable in guiding future CDS implementations.

Finally, an important implication of our findings is that meta-regression analyses can potentially lead to erroneous conclusions. Therefore, caution should be exercised in making decisions based on the findings of meta-regression analyses alone.

4.3. Recommendations

Given the perils of meta-regression as described in this case study, we recommend the following moving forward:

- Support future meta-analyses by clearly identifying key features of CDS systems in published evaluations in order to facilitate effective explanatory factor abstraction.
- Give careful consideration to the definition and abstraction of potential explanatory factors when conducting meta-regression analyses.
- Consider meta-regression as an interim step for generating hypotheses for the importance of specific health information system features. Any findings should be validated through actual implementation experience and clinical trials.
- Identify mechanisms for promoting reanalyses of meta-regression studies. For example, a journal may devote a special issue focused on replication studies, including for meta-regression analyses.
- Conduct more trials designed to directly evaluate the importance of specific CDS features. Once available, generally defer to the results of direct trials rather than regression analyses.

5. Conclusion

Given the limitations of meta-regression analyses, further research – particularly head to head trials evaluating the effectiveness of CDS system features – is clearly warranted.

Funding

We would like to acknowledge the University of Utah Department of Biomedical Informatics and the Intermountain Healthcare Homer Warner Center for Informatics Research for sponsoring this research.

This study was funded in part by National Library of Medicine training Grant T15-LM07124 (CLF), the University of Utah Program in Personalized Health Care (BMW), and the University of Utah Department of Biomedical Informatics (KK, MZ, and CAR).

Ethical approval

None required.

Data sharing

Statistical code and the dataset are available from the corresponding author.

Conflict of interest

KK is currently or recently served as a consultant on CDS to the Office of the National Coordinator for Health IT, ARUP Laboratories, McKesson InterQual, ESAC, Inc., JBS International, Inc., Inflexxion, Inc., Intelligent Automation, Inc., Partners HealthCare, Mayo Clinic, and the RAND Corporation. KK receives royalties for a Duke University-owned CDS technology for infectious disease management known as CustomID that he helped develop. KK was formerly a consultant for Religent, Inc. and a co-owner and consultant for Clinica Software, Inc., both of which provide commercial CDS services, including through use of a CDS technology known as SEBASTIAN that KK developed. KK no longer has a financial relationship with either Religent or Clinica Software. All other authors declare no competing interests.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jbi.2015.05.007.

References

[1] K. Kawamoto, C.A. Houlihan, E.A. Balas, et al., Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success, BMJ 330 (7494) (2005) 765.
[2] A.X. Garg, N.K. Adhikari, H. McDonald, et al., Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review, JAMA 293 (10) (2005) 1223–1238.
[3] K. Kawamoto, D.F. Lobach, Clinical decision support provided within physician order entry systems: a systematic review of features effective for changing clinician behavior, AMIA Annu Symp Proc (2003) 361–365.
[4] D. Lobach, G.D. Sanders, T.J. Bright, et al., Enabling health care decisionmaking through clinical decision support and knowledge management, Evidence Report/Technology Assessment 203 (2012) 1–784.
[5] E.G. Poon, A. Wright, S.R. Simon, et al., Relationship between use of electronic health record features and health care quality: results of a statewide survey, Med. Care 48 (3) (2010) 203–209.
[6] P.S. Roshanov, N. Fernandes, J.M. Wilczynski, et al., Features of effective computerised clinical decision support systems: meta-regression of 162 randomised trials, BMJ 346 (2013) f657.
[7] J.P. Higgins, S.G. Thompson, Controlling the risk of spurious findings from meta-regression, Stat. Med. 23 (11) (2004) 1663–1682.
[8] S.G. Thompson, J.P. Higgins, How should meta-regression analyses be undertaken and interpreted?, Stat Med. 21 (11) (2002) 1559–1573.
[9] H.C. van Houwelingen, L.R. Arends, T. Stijnen, Advanced methods in meta-analysis: multivariate approach and meta-regression, Stat. Med. 21 (4) (2002) 589–624.
[10] T.J. Bright, A. Wong, R. Dhurjati, et al., Effect of clinical decision-support systems: a systematic review, Ann. Int. Med. 157 (1) (2012) 29–43.
[11] G. Damiani, L. Pinnarelli, S.C. Colosimo, et al., The effectiveness of computerized clinical guidelines in the process of care: a systematic review, BMC Health Services Research 10 (2010) 2.
[12] D. Dorr, L.M. Bonner, A.N. Cohen, et al., Informatics systems to promote improved care for chronic illness: a literature review, J. Am. Med. Inform. Assoc. 14 (2) (2007) 156–163.
[13] A.M. Scheepers-Hoeks, R.J. Grouls, C. Neef, et al., Physicians' responses to clinical decision support on an intensive care unit-comparison of four different alerting methods, Artif. Intell. Med. 59 (1) (2013) 33–38.
[14] J.T. Van Wyk, M.A.M. Van Wijk, M.C.J.M. Sturkenboom, et al., Electronic alerts versus on-demand decision support to improve dyslipidemia treatment: a cluster randomized controlled trial, Circulation 117 (3) (2008) 371–378.
[15] Stata Statistical Software: reLease 13 [program]. College Station, TX, StataCorps, 2014.
[16] D.A. Fitzmaurice, F.D.R. Hobbs, E.T. Murray, et al., Oral anticoagulation management in primary care with the use of computerized decision support and near-patient testing: a randomized, controlled trial, Arch. Int. Med. 160 (15) (2000) 2343–2348.
[17] A. Boguševičius, A. Maleckas, J. Pundzius, et al., Prospective randomised trial of computer-aided diagnosis and contrast radiography in acute small bowel obstruction, Eur. J. Surg. 168 (2) (2002) 78–83.